LEARNING DATA AUGMENTATION POLICIES

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit under 35 U.S.C. § 119(e) of U.S. Patent Application No. 62/673,777, entitled "LEARNED DATA AUGMENTATION," filed May 18, 2018. The disclosure of the foregoing application is incorporated herein by reference in its entirety for all purposes.

BACKGROUND

This specification relates to processing data using machine learning models.

Machine learning models receive an input and generate an output, e.g., a predicted output, based on the received input. Some machine learning models are parametric models and generate the output based on the received input and on values of the parameters of the model.

Some machine learning models are deep models that employ multiple layers of models to generate an output for a received input. For example, a deep neural network is a deep machine learning model that includes an output layer and one or more hidden layers that each apply a non-linear transformation to a received input to generate an output.

SUMMARY

This specification describes a training system implemented as computer programs on one or more computers in one or more locations.

According to a first aspect there is provided a method that includes receiving training data for training a machine learning model to perform a particular machine learning task. The training data includes multiple training inputs. Multiple data augmentation policies are determined, where each data augmentation policy has multiple data augmentation policy parameters that define a procedure for transforming training inputs before the training inputs are used to train a machine learning model.

At each of multiple time steps, a current data augmentation policy is generated based on quality measures of data augmentation policies generated at previous time steps. A quality measure of a data augmentation policy represents a performance of a machine learning model on a particular machine learning task as a result of training the machine learning model using the data augmentation policy. A machine learning model is trained on the training data using the current data augmentation policy. Training a machine learning model using a data augmentation policy includes: selecting a batch of training data, determining an augmented batch of training data by transforming the training inputs in the batch of training data in accordance with the data augmentation policy, and adjusting current values of the machine learning model parameters based on the augmented batch of training data. A quality measure of the current data augmentation policy is determined using the machine learning model after it has been trained using the current data augmentation policy.

A final data augmentation policy is selected based on the quality measures of the determined data augmentation policies. A final trained machine learning model is generated by training a final machine learning model using the final data augmentation policy.

In some implementations, the particular machine learning task is an image processing task including classification or regression.

In some implementations, each data augmentation policy includes one or more sub-policies. Each sub-policy includes a sequence of one or more transformation tuples, where for each transformation tuple, the data augmentation policy parameters define: (i) a transformation operation, and (ii) a magnitude of the transformation operation. Transforming the training inputs in the batch of training data in accordance with the data augmentation policy includes, for each training input: identifying a sub-policy included in the data augmentation policy; and transforming the training input in accordance with the identified sub-policy by sequentially applying each transformation tuple included in the identified sub-policy to the training input.

In some implementations, identifying a sub-policy included in the data augmentation policy for the training input includes randomly sampling a sub-policy included in the data augmentation policy.

In some implementations, applying a transformation tuple to the training input includes applying the transformation operation from the transformation tuple with the transformation operation magnitude from the transformation tuple to the training input.

In some implementations, for each transformation tuple, the data augmentation policy parameters further define a probability of applying the transformation operation; and applying a transformation tuple to the training input includes applying the transformation operation from the transformation tuple with the transformation operation magnitude from the transformation tuple to the training input with the transformation probability from the transformation tuple.

In some implementations, the machine learning model is a neural network, and adjusting the current values of the machine learning model parameters based on the augmented batch of training data includes: determining a gradient of a loss function using the augmented batch of training data; and adjusting the current values of the machine learning model parameters using the gradient.

In some implementations, generating the current data augmentation policy based on quality measures of data augmentation policies generated at previous time steps includes generating the current data augmentation policy using a policy neural network in accordance with current values of policy neural network parameters; and the policy neural network is trained by reinforcement learning techniques, and at each time step, the reinforcement learning reward signal is based on the quality measure of the current data augmentation policy at the time step.

In some implementations, for each data augmentation policy parameter, the policy network output defines a score distribution over possible values of the data augmentation policy parameter.

In some implementations, determining the current data augmentation policy from the policy network output includes sampling a value for each data augmentation policy parameter using the score distribution for the data augmentation policy parameter.

In some implementations, the policy neural network is a recurrent neural network.

In some implementations, generating a current data augmentation policy based on quality measures of data augmentation policies generated at previous time steps includes generating the current data augmentation policy using a genetic programming procedure.